Audio inpainting: problem statement, relation with sparse representations and some experiments

Amir Adler, Valentin Emiya, Maria G. Jafari, Michael Elad, Rémi Gribonval, Mark D. Plumbley

Audio inpainting: problem statement, relation with sparse representations and some experiments

Amir Adler[1], **Valentin Emiya**[2], **Maria Jafari**[3], **Michael Elad**[1], **Rémi Gribonval**[2], **Mark Plumbley**[3]

[1] The Technion, Israel [2] INRIA Rennes - Bretagne Atlantique, France [3] Queen Mary University of London, UK

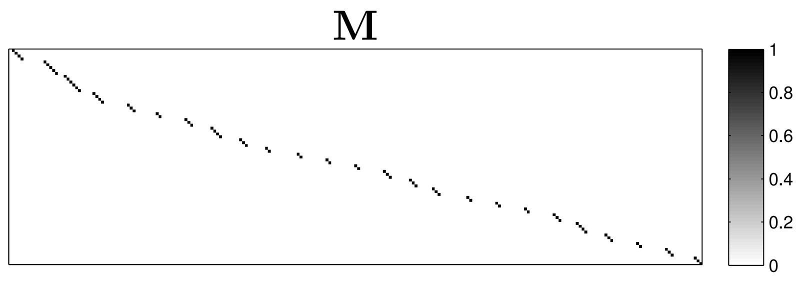
Problem statement & applications: a unified scheme for existing tasks

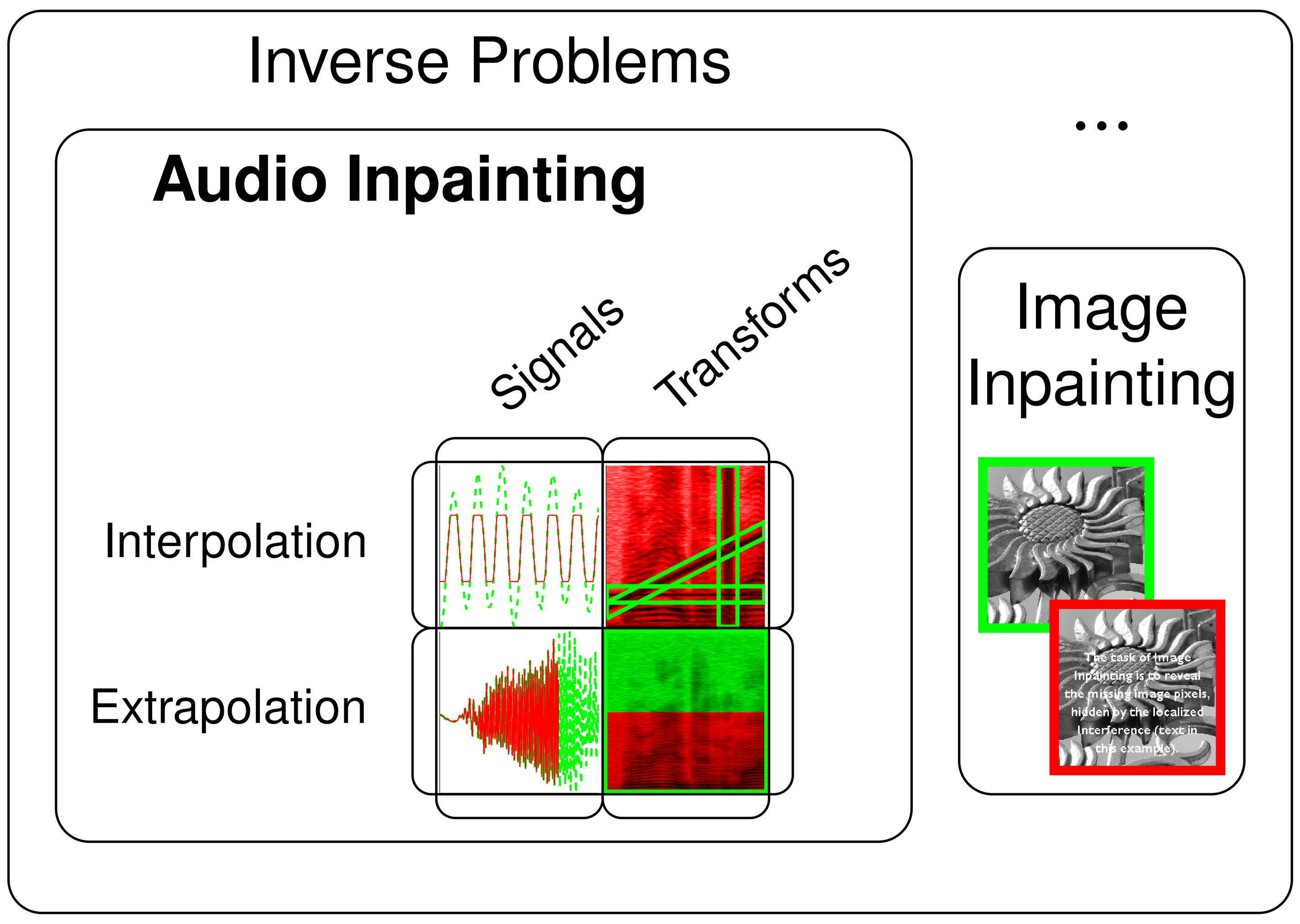
Audio inpainting scheme: an inverse problem where
- a set of reliable audio data $\mathbf{y}$ is observed,
- one must estimate the remaining missing or highly corrupted data.

→ A unified scheme covering existing subproblems referred to as interpolation, extrapolation, imputation, (bandwidth) extension.

Formulation: estimate the original data $\mathbf{s}$ from $\mathbf{y}$ given M

(1) $\mathbf{y} = \mathrm{M}\mathbf{s}$ with $\begin{cases} \mathbf{s} & \in \mathbb{R}^N \\ \mathbf{y} & \in \mathbb{R}^{N-M} \\ \mathrm{M} & \in \mathbb{R}^{N-M} \times \mathbb{R}^N \end{cases}$ and





Several kinds of audio data: waveforms [1,2], transforms or mid-level representations [3].

A number of applications: removing clicks in old recordings, declipping, packet loss concealment in VoIP or P2P networks, bandwidth extension, recovery of TF coefficients masked by interfering sources/noise.

Proposed approach: audio inpainting using sparse representations

Since the problem (1) is ill-posed, additional *a priori* is required.

Sparsity assumption on audio data

(2) $\mathbf{s} = \mathbf{D}\mathbf{x}$ with $\begin{cases} \mathbf{D} \in \mathbb{R}^N \times \mathbb{R}^{K_D} \text{ (dictionary)} \\ \mathbf{x} \in \mathbb{R}^{K_D} \text{ (sparse coefficients)} \\ \|\mathbf{x}\|_0 \ll N \leq K_D \end{cases}$

Noiseless ideal estimation

(3) $\hat{\mathbf{x}} \triangleq \arg\min_{\mathbf{x}} \|\mathbf{x}\|_0$ s.t. $\mathbf{y} = \mathrm{M}\mathbf{D}\mathbf{x}$

Inpainting algorithms

Algorithm 1 Inpainting with l_1-minimization (L_1)

Using convex minimization, do

$$\hat{\mathbf{x}} \leftarrow \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ s.t. } \|\mathbf{y} - \mathrm{M}\mathbf{D}\mathbf{x}\|_2^2 \leq \theta^{l_2}$$

$\hat{\mathbf{s}} \leftarrow \mathbf{D}\hat{\mathbf{x}}$

Algorithm 2 OMP-based inpainting (OMP)

Dictionary $\widetilde{\mathbf{D}} = \left[\widetilde{\mathbf{d}}_1, \ldots, \widetilde{\mathbf{d}}_{K_D}\right] \leftarrow \mathrm{M} \times \mathbf{D} \times \mathbf{W}_{\mathrm{MD}}^{-1}$ ($\mathbf{W}_{\mathrm{MD}}$: diag. matrix of norms of MD columns)
Residual $\mathbf{r}_0 \leftarrow \mathbf{y}$
Iteration counter $k \leftarrow 1$, support set $\Omega_0 \leftarrow \emptyset$
while $k \leq K_{max}$ AND $\|\mathbf{r}_k\|_2 \geq \theta_{OMP}$ **do**
  Select atom: $j \leftarrow \arg\max_j |<\mathbf{r}_{k-1}, \widetilde{\mathbf{d}}_j>|$
  Update support $\Omega_k \leftarrow \Omega_{k-1} \cup j$
  Update current solution $\mathbf{x}_k \leftarrow \arg\min_{\mathbf{x}} \|\mathbf{y} - \widetilde{\mathbf{D}}_{\Omega_k}\mathbf{x}\|_2$
  Update residual $\mathbf{r}_k \leftarrow \mathbf{y} - \widetilde{\mathbf{D}}_{\Omega_k}\mathbf{x}_k$
  Increment iteration counter $k \leftarrow k + 1$
end while
$\hat{\mathbf{s}} \leftarrow \mathbf{D}\mathbf{W}_{\mathrm{MD}}\mathbf{x}_k$

Specific inpainting algorithms for restoring clipped signals

→ Constrain the restored samples to be beyond the clipping threshold θ_{clip}

Algorithm 3 (L_1c)

Using convex minimization, do

$$\hat{\mathbf{x}} \leftarrow \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ s.t. } \begin{cases} \|\mathbf{y} - \mathrm{M}\mathbf{D}\mathbf{x}\|_2^2 \leq \theta^{l_2} \\ \mathrm{M}^+\mathbf{D}\mathbf{x} \geq \theta_{\text{clip}} \\ \mathrm{M}^-\mathbf{D}\mathbf{x} \leq -\theta_{\text{clip}} \end{cases}$$

$\hat{\mathbf{s}} \leftarrow \mathbf{D}\hat{\mathbf{x}}$

Algorithm 4 (OMPc)

After the while loop in Alg. 1, refine $\mathbf{x}_k$ as

$$\mathbf{x}_k \leftarrow \arg\min_{\mathbf{x}} \|\mathbf{y} - \widetilde{\mathbf{D}}_{\Omega_k}\mathbf{x}\|_2 \text{ s.t. } \begin{cases} \mathrm{M}^+\mathbf{D}\mathbf{x} \geq \theta_{\text{clip}} \\ \mathrm{M}^-\mathbf{D}\mathbf{x} \leq -\theta_{\text{clip}} \end{cases}$$

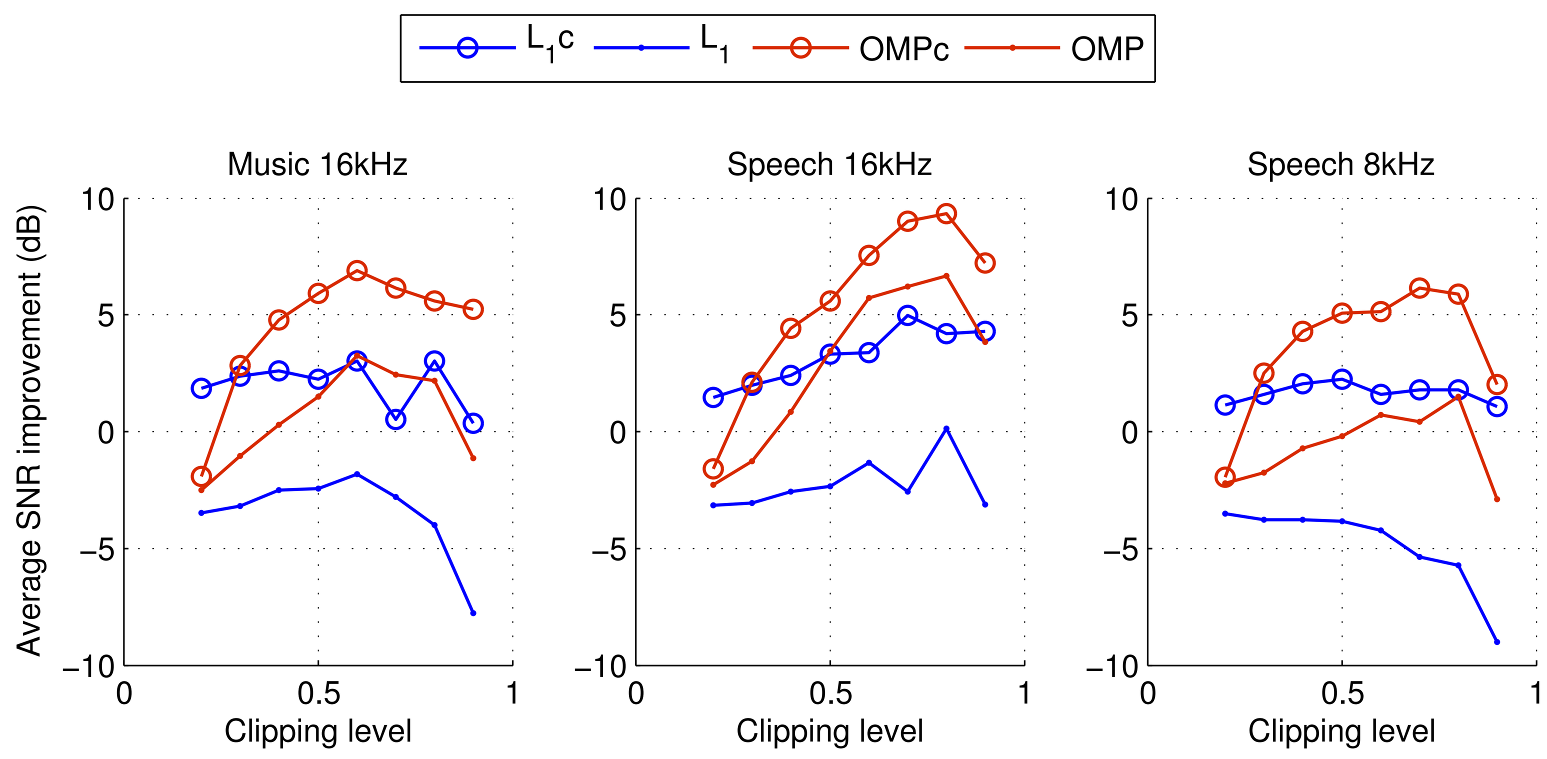
$\hat{\mathbf{s}} \leftarrow \mathbf{D}\mathbf{W}_{\mathrm{MD}}\mathbf{x}_k$

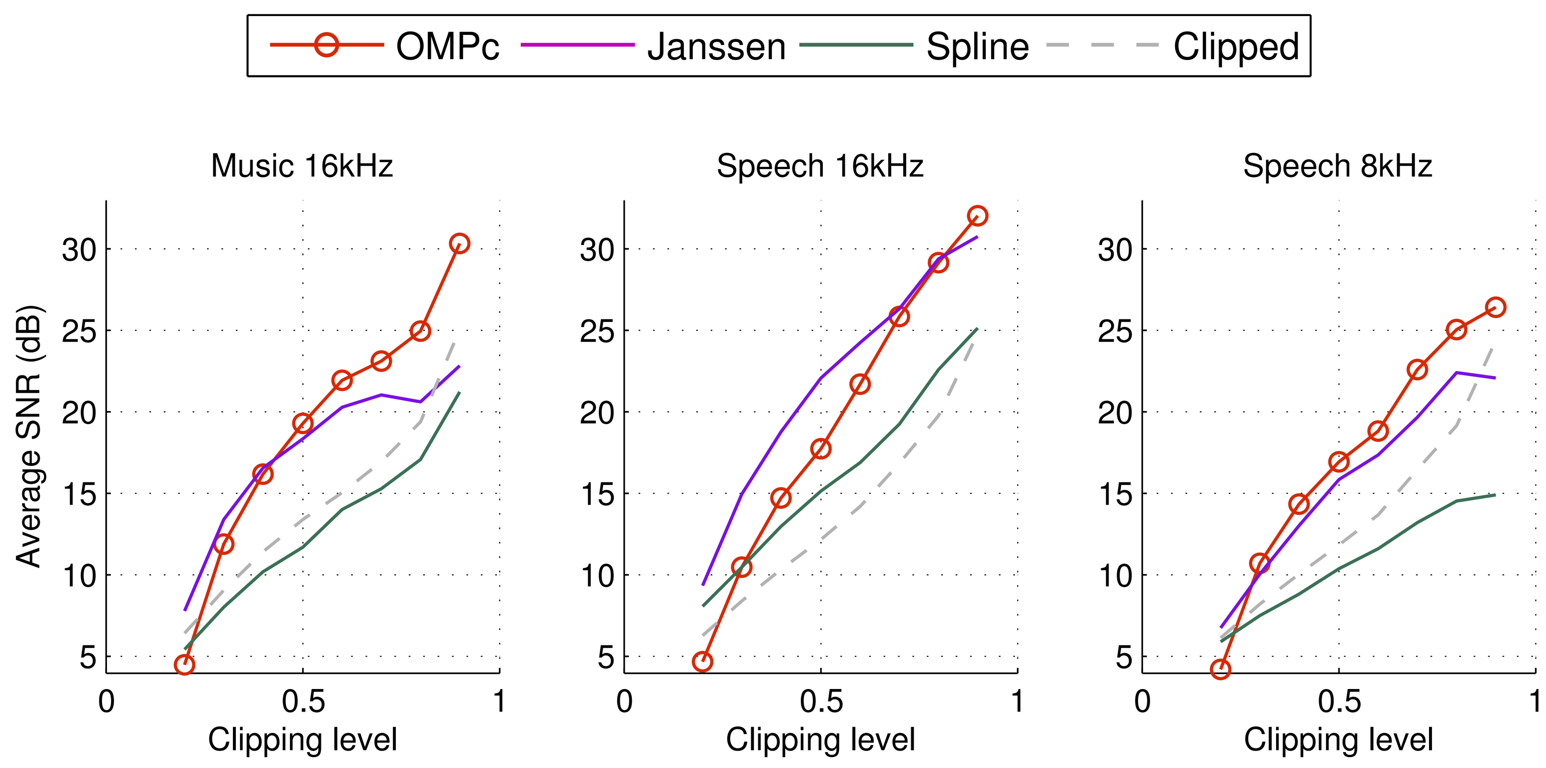
where M^+ and M^- are the projectors onto the subspace of positive and negative clipped samples respectively.

Experiments: declipping audio signals

Experimental settings

- 3 datasets with 10 5-seconds signals
- Frame-by-frame processing with OLA reconstruction: 75% overlap, 64ms frames, sine weighting windows
- Algorithms OMP, OMPc, L_1, L_1c are applied in each frame
- OMPc is compared to spline interpolation and Janssen's algorithm [1] based on autoregressive models.
- Algorithm parameters are fixed (tuned on a separate database)
- Dictionary $\mathbf{D}$: redundant sine-windowed DCT
- SNR is computed on the corrupted samples only





[1] A. Janssen, R. Veldhuis, L. Vries, *Adaptive interpolation of discrete-time signals that can be modeled as autoregressive processes*, in IEEE Trans. on Acoustics, Speech and Signal Proc., 34 (2), 1986.
[2] S. J. Godsill, P. J. W. Rayner, *Digital audio restoration - A statistical model-based approach*, Springer-Verlag London, 1998.
[3] J. Le Roux, H. Kameoka, N. Ono, A. de Cheveigné, S. Sagayama, *Computational auditory induction by missing-data non-negative matrix factorization*, Proc. of SAPA, 2008.

9th Int. Conf. on Latent Variable Analysis and Signal Separation, September 27-30, 2010, St. Malo, France

This work was supported by the European Union through the project SMALL, under FET-Open grant agreement no. 225913.